Image classification of sugar crystal with deep learning

Suriya Chayatummagoon
Computer Engineering Department,
Faculty of Engineering
Chulalongkorn University
Bangkok, Thailand

Prabhas Chongstitvatana
Computer Engineering Department,
Faculty of Engineering
Chulalongkorn University
Bangkok, Thailand

Abstract— **Nowadays, artificial intelligent control is essential to replace experienced workers. The correct classification of sugar crystals during the production process is the basis for the control of the sugar crystallization process. Correct Classification of sugar crystals is the basis necessary for automatic control of process. This research uses the principles of deep learning using a neural network to identify the crystallization of sugar from the actual production process of sugar factories in Thailand. Performance was measured and compared with the Fine-tuning VGG16 model. It was accurate to identify sugar crystals between 82% and 92% of four classes sugar crystal images classified by the crystallization conditions. The results of this study also show that this model is more accurate than other models. It can be used as a benchmark for monitoring the crystallization of sugar production processes. It is also the basis of an artificial intelligent control system based on transcribing human expertise**

Keywords-component; Deep learning, Image Classification, Crystal Formation

I. INTRODUCTION

The crystallization process still requires experienced sugar pickers. For controlling the crystallization process to have quality but the practice of sugar simmering requires a long time. Each person has different abilities and expertise. Therefore, it affects the productivity of the plant. In addition, the wages of an experienced sugar cook are higher. When compared with other process controllers an experienced sugar cooker will be demanded by the labor market in this industry causing sugar mills to always face a shortage of personnel in this position also inspired by the sugar crystallization control process was being tested in Japan in early 2020 using digital image processing [1,2]. Crystalline images are continuously sent to the computer in the control room in which specific software applies a specific algorithm to each image and calculates including the coefficient of variation and the mean of the space between the crystals ratio is the percentage of a small, non-crystalline entity. The number of tablets in real time adjustable criteria for variables are used to detect and warn of inconsistencies such as the presence of contaminants at a given time.

A. *Sugar Crystal Theory*

Boiling syrup and crystallization is an important process that can be considered as the primary process for transforming sugar into a liquid syrup solution and then crystalline sugar into solid crystalline form [3, 4]. The general principles of the crystallization process require to extract the latent heat of the crystalline solid out of the solution to be crystallized to cause a liquid to solidify at the temperature of that crystallization. Therefore, in the process the relationship of the mass distribution in the liquid and solid phases must be known. The process of raising crystals to meet the standard requirements in this regard, both the boiling and chewing process must be controlled. The process of raising crystals and the process of crystallization to get the maximum amount of sugar crystals however, require boiling, simmering and crystallization. They are more difficult to control than normal conditions. This is because the liquid involved in this process is in the form of a very concentrated syrup. With a very high viscosity, it is difficult to control the flow or the flow through the pipe system. The heat transfer in the system must also be considered. The measure of supersaturated concentration indicates three distinct state of the crystals. From 1.00 - 1.20 is called metastable zone, between 1.20 - 1.30 is called intermediate zone and from 1.30 - 1.40 is called labile zone. The range of these values can be applied to control the sugar crystallization process depending on the technique of each plant. Each phase of this supersaturated concentration has a characteristic syrup as shown in Fig 1.

Metastable zone is a time when the syrup is slightly over saturation point, there will be an excess of sugar molecules in the saturation point. But it is not yet dense enough for the sugar molecules to bind together to form nuclei.

Intermediate zone is the period when the syrup has a higher density of sugar molecules that allow some point of the syrup to form spontaneous nuclei and will gradually become larger. While the crystals are formed, new nuclei may be produced, causing the resulting crystals to be of different sizes.

Labile zone is the period when the syrup is so concentrated that the number of sucrose molecules is so dense and close together that it produces many free nuclei. During the super-saturated concentration and when the syrup solution was so viscous that the crystallization was very difficult or might not be able to control at all.

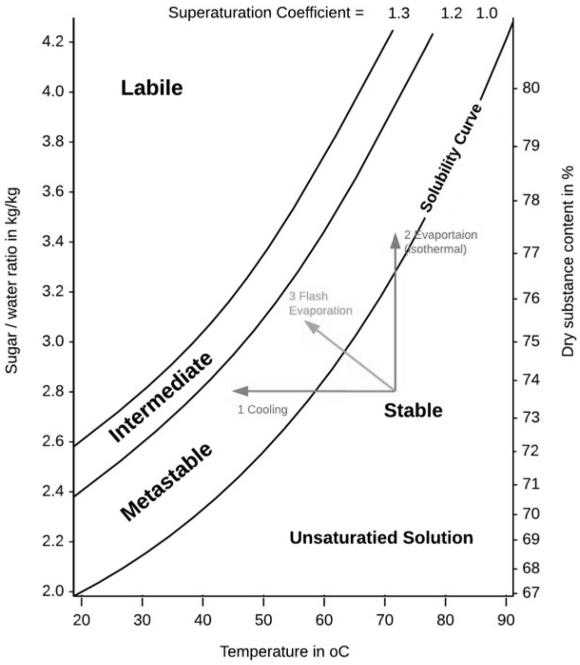

Fig. 1. Relationship of the state of sugar crystallization.

B. *Deep learning theory*

Deep learning is the concept of a computer learning to understand the information provided [5]. Convolutional neural networks [6] are used to extract distinctive features and to realize patterns from photographs. Convolutional neural networks are a type of feed forward process. The convolutional layers are used to filter images with a kernel function to convert features and to separate elements, such as border, color, shape, etc. A layer called pooling serves to reduce the size of the data. The last part is a fully connected layer that will be used to make decision, Fig 2.

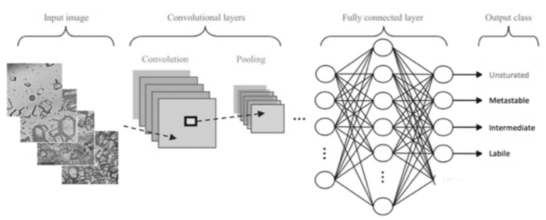

Fig. 2. CNN. of crystal prediction, output 4 classes.

II. RELATED RESEARCH

A number of works has been done on classification of crystallization outcomes using deep convolutional neural networks. The work in [6] uses crystal images from five data sources with nearly half a million images for classification. It was found that more than 94% of the test images were able to accurately predict the type of crystal. The work from sugar factories in China [2] uses 1,000 crystals images. Three experienced humans were asked to divide the sugar crystal classes into 5 classes. With this label data, they experiment with 4 transfer learning models. The work in [7] carried out at a factory in Spain. They use spectral images with neural networks to classify sugar crystallization. Their work divided the output into 4 classes (two sizes, deformed crystals, homogeneous, three or more size).

III. EQUIPMENT AND METHODS

Sampling data of sugar crystals were collected from the simmering pot during the simmering process. By recording sugar crystals on a clear glass pretzel shines through x10 magnification glasses with a rear light source, see Fig. 3. The crystallization process of cane sugar is a process in which sucrose absorbs nutrients from the syrup. Sucrose crystals are colorless and transparent. They are slightly separated from the background. This allows computers to accurately distinguish between the target and background during the recognition process.

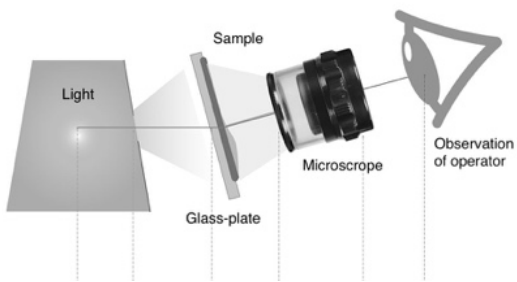

Fig. 3. Camera with x10 microscope.

A. *DATA COLLECTION*

The images are captured from the location at single sight glass. This is the reference position which is used in [1]. The resulting images have the same quality as the observation of the simmering staff. It is the same view this is used to make decision in order to control the browning process. Fig. 4 shows crystals in various zones

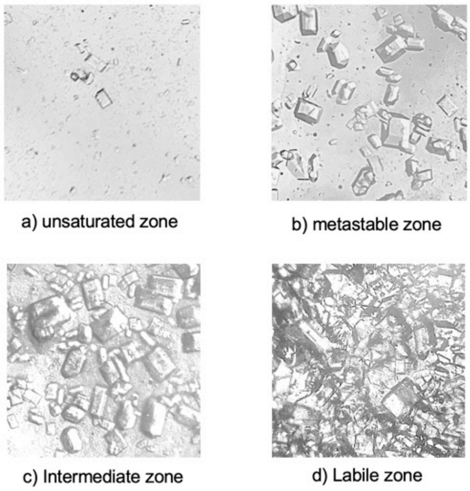

Fig. 4. Images of sugar crystal in various zones

The total number of sugar crystal images used in the data set is 1,000 images. They are divided into 800 Training set and 200 Testing set. The crystal images were then separated by experienced stewers who divide the crystals into 4 classes: Unsaturated Zone, Metastable Zone, Intermediate Zone and Labile Zone. The number of crystals collected in each zone is 250 images per zone as shown in Table I.

TABLE I. CLASS DATA OF THE TYPE OF CRYSTALS

Class	No. Images	
	Training Dataset	Testing Dataset
Unsaturated Zone	200	50
Metastable Zone	200	50
Intermediate Zone	200	50
Labile Zone	200	50

The validation of the recorded crystal images was done by concluding opinions from 5 skilled stewers. The verification process of the crystal photographs is done by making a test for 32 images per one employee with the same number of random crystal images tested. This test has an average accuracy of 95% and the value is between 92 - 98%.

B. Data preparation

The images are preprocessed with 5 operations. The results are shown in Fig. 5.

- Image adjustment to be black and white
- Random brightness adjustment (± 30%)
- Random saturation adjustments
- Random chroma adjustments
- Random sharpening (from 50% to 150%)

To increase the variety of the training data, the images are inverted with 50% probability, then rotate the image at 90˚, 180˚, 270˚ degrees with 50% probability, and then crop the image in the center 150 x 150 pixels.

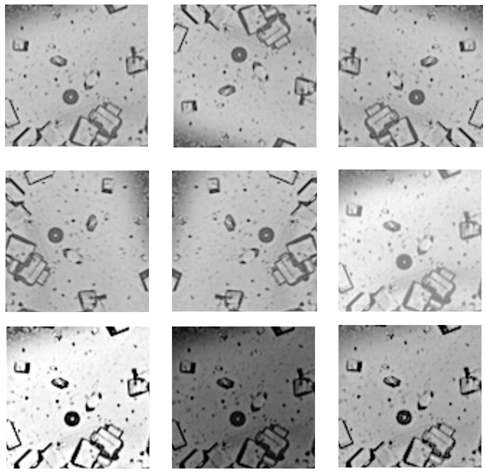

Fig. 5. Crystal images after processing

C. Simple Model

The software components consisted of python 3.8, tensorflow 2.3, keras 2.4.3 on the kaggle twinTPU v3.8 16GB.

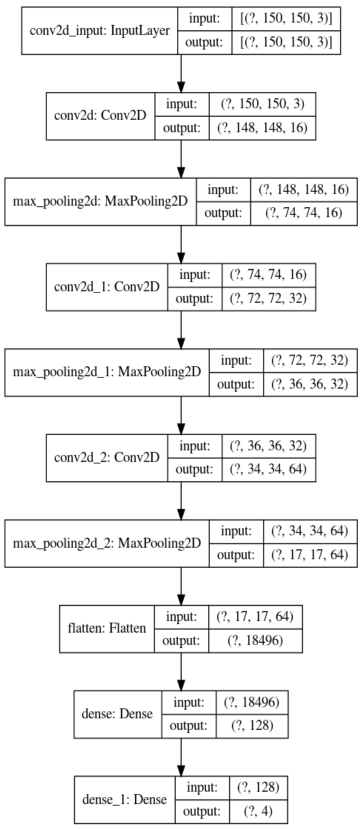

Fig. 6. Structure of model simple DCNNs.

The training dataset of 800 images with 4 classes (200 images/class). The first model of the experiment is a simple DCNNs. The following parameters are used: convolution2d, maxpooling2d with 3 step 16, 32 and 64 filters having dimensions of 3 x 3 and dense 128 and 4 using adam optimizer with learning rate 0.0001. The model is fitted with batch size 128 training, 15 epochs. Fig 6 shows the structure of the simple DCNNs model.

D. Transfer Learning

The second experiment employs transfer learning [8]. Training the complex deep learning model from scratch (weight initialization with random values) requires both a large training data and high computational power. Therefore, using transfer learning is a popular technique because it reduces the training time by incorporating some of the already trained models with similar tasks as part of the new model.

Three models are used in the experiment: inception-v3, ResNet50-v2 and VGG16 [6]. They are compared to DCNNs models. All 3 models load weight from imagenet and custom head by adding flatten layer.

The performance of all four models are compared and the results are shown in Table II with accuracy, precision and recall measurements.

TABLE II. Comparing four models

Model	Accuracy	Precision	Recall
Simple DCNNs	50.5	54.0	50.5
Inception-v3-tf	78.5	78.5	78.5
ResNet50-v2-tf	77.5	82.5	77.5
VGG16-tf	85.0	87.4	85.0

From Table II, the best result came from VGG16-tf. It has highest accuracy, precision and recall.

E. Fine-Tuning

The model VGG16 transfer learning has been so effective on the first step. To improve the performance furthur some parameters in the transfer learning are adjusted to find the optimal values. There are 3 experiments 1) using 10 folds [9] by split the dataset and train 10 cycles to reduce overfitting, 2) customize lasted 5 layers with fixed weight in base model and 3) customize lasted 5 layer unfixed weight in base model. By customizing this layer, it will convert the original 5 layers of VGG16 to conv2d 3x3 pixels size with 64 filters (a), maxpooling2d size 2x2 pixel (b), flatten layer (c), dense 100 using relu activation (d). The output consists of dense 4 using softmax activation, using adam optimizer learning rate 0.0001 and fit model in batch size 128 training, 15 epochs. The results of the experiments were measured using accuracy, precision and recall are shown in Table III.

From Table III, the VGG16 fine-tuning model (c) with custom lasted 5 layers with fixed weight in base model in first 15 layers, has the highest accuracy. Fig. 7 shows the confusion matrix [10] of VGG16-tf model c.

TABLE III. Tuning the model performance

Fine-tuning VGG16 models	Accuracy	Precision	Recall
a) Custom head	85.0	87.4	85.0
b) Estimator 10 folds	83.0	86.0	83.0
c) Custom last 5 fixed weight	89.0	90.0	89.0
d) Custom last 5 fixed non-weight	87.5	88.0	87.5

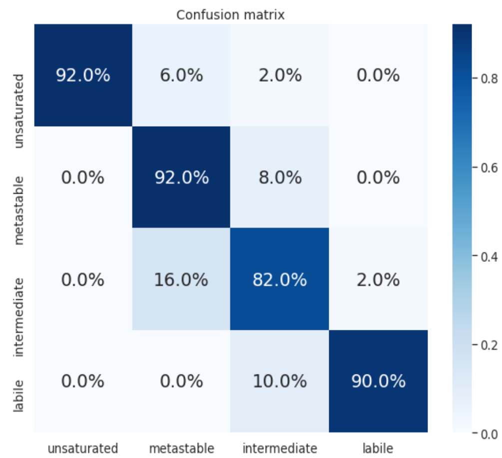

Fig 6. Confusion matrix VGG16 fine tuning model c

IV. Results

In the experiments, the prediction accuracy of Class 3 is 82%, Class 4 is 90%, and Class1 and 2 are of 92%. Sugars in the saturated (Class1) range early in the simmering test were different from the metastable (Class 2) range, resulting in higher prediction compared to intermediate (Class 3) with the lowest prediction value from the crystalline form. This range is similar to the labile (Class 4) and such simmering, the steam conditions that affect the simmering sugar during the maintenance of the main boiler.

V. Conclusion

In this research, deep learning was applied to investigate the crystallization of sugar. The data are collected from the packing season 2019/2020 of sugar factories in Thailand. The model is based on visual data of the crystallization of sugar from the actual production process. The expertise of five experienced simmerers are used to validate the data set.

Through the vector feature analysis in the transfer learning process, the model with the highest accuracy and accuracy is Model C, custom 5 last layer by fixed weight in base model VGG16. The prediction accuracy of all four classes is range between 82% to 92%. This model can be deployed in a docker environment to implement the automatic process control [11].

REFERENCES

[1] A. Iteca, “New Process Control During The Crystal Growth: Real Time Image Analysis, Comparison of Strikes,” Sugar Asia, Jan 2020.

[2] J. Zhang, Y. Meng, J. Wu and S. Yu, “Monitoring sugar crystallization with deep neural networks,” Journal of Food Engineering,Volume 280,109965, Feb 2020.

[3] G. Jenkins, “Introduction to Cane Sugar Technology,” Science Direct, pp.267-375, 1966.

[4] H. Pieter,“Principles of Sugar Technology,” Bibliography of Crystallized, Elsevier, pp. 234-257.

[5] F. Ertam and G. Aydin, “Data classification with deep learning using Tensorflow”, 2017 Int Conf on Computer Science and Engineering, pp. 755-758.

[6] A. Bruno, P. Charbonneau, J.Newman, E. Snell and D. So, “Classification of crystallization outcomes using deep convolutional neural networks,” PLoS ONE 13(6): e0198883, June 2018.

[7] B.G. Mertzios and P. Liatsis,“Techniques for classifying sugar crystallization images based on spectral analysis and the use of neural networks,” IWISP '96, Elsevier, pp.641-645, Nov 1996.

[8] S. R. Dhruba, R. Rahmanl, K. Matlockl, S. Ghosh and R. Pal, “Dimensionality Reduction based Transfer Learning applied to Pharmacogenomics Databases,” Int Conf of the IEEE Engineering in Medicine and Biology Society, 2018, pp. 1246-1249, doi: 10.1109/EMBC.2018.8512457.

[9] S. Yadav and S. Shukla, “Analysis of k-Fold Cross-Validation over Hold-Out Validation on Colossal Datasets for Quality Classification,” IEEE Int Conf on Advanced Computing, 2016, pp. 78-83, doi: 10.1109/IACC.2016.25.

[10] N. D. Marom, L. Rokach and A. Shmilovici, “Using the confusion matrix for improving ensemble classifiers,” 2010 IEEE Convention of Electrical and Electronics Engineers in Israel, 2010, pp.555-559, doi: 10.1109/EEEI.2010.5662159.

[11] L. Suarez, P. Georgieva, S.F. de Azevedo, “Computationally efficient process control with neural networkbased predictive models,” Int Joint Conf on Neural Networks, 2009, pp.2990-2997.